

Dependence of multivariate extremes

C. Viseu *

Instituto Superior de Contabilidade
e Administração de Coimbra
Portugal

L. Pereira, A.P. Martins, H. Ferreira †

Departamento de Matemática
Universidade da Beira Interior
Portugal

Abstract: We give necessary and sufficient conditions for two sub-vectors of a random vector with a multivariate extreme value distribution, corresponding to the limit distribution of the maximum of a multidimensional stationary sequence with extremal index, to be independent or totally dependent. Those conditions involve first relations between the multivariate extremal indexes of the sequences and secondly a coefficient that measure the strength of dependence between both sub-vectors. The main results are illustrated with an auto-regressive sequence and a 3-dependent sequence.

1 Introduction

Multivariate Extreme Value Analysis is frequently applied in the context of modeling environmental data, for which the phenomenon of dependence is often intrinsic. This paper focuses on the characterization of total dependence and of independence of two multivariate extreme value distributions.

Let $\mathbf{X} = \{\mathbf{X}_n^{(d)} = (X_{n,1}, \dots, X_{n,d})\}_{n \geq 1}$ be a d -dimensional stationary sequence with common distribution function (d.f.) $Q(\mathbf{x}^{(d)}) = Q(x_1, \dots, x_d)$, $\mathbf{x}^{(d)} \in \mathbb{R}^d$, and $\mathbf{M}_n = (M_{n,1}, \dots, M_{n,d})$ the vector of pointwise maxima, where $M_{n,i}$ is the maximum of i -th

*Instituto Superior de Contabilidade e Administração de Coimbra, Coimbra, Portugal.

E-mail: cviseu@iscac.pt

†Departamento de Matemática, Universidade da Beira Interior, 6200 Covilhã, Portugal. E-mail: lperreira@ubi.pt; amartins@ubi.pt; helena.ferreira@ubi.pt

component of \mathbf{X} . Denote $\widehat{\mathbf{M}}_n = (\widehat{M}_{n,1}, \dots, \widehat{M}_{n,d})$ the corresponding vector of pointwise maxima of the associated d -dimensional sequence, $\widehat{\mathbf{X}} = \{\widehat{\mathbf{X}}_n^{(d)}\}_{n \geq 1}$, of independent and identically distributed (i.i.d.) random vectors having the same distribution function Q .

In this multivariate setting operations among vectors are defined componentwise, that is for each $d > 1$ and $\mathbf{a}^{(d)}, \mathbf{b}^{(d)} \in \mathbb{R}^d$, $\mathbf{a}^{(d)} \leq \mathbf{b}^{(d)}$, if and only if $a_j \leq b_j$, for all $j = 1, 2, \dots, d$.

If there exist sequences $\{(a_{n,1} > 0, \dots, a_{n,d} > 0)\}_{n \geq 1}$ and $\{(b_{n,1}, \dots, b_{n,d})\}_{n \geq 1}$, such that for $\mathbf{u}(\mathbf{x}^{(d)}) = \{\mathbf{u}_n(\mathbf{x}^{(d)}) = (a_{n,1}x_1 + b_{n,1}, \dots, a_{n,d}x_d + b_{n,d})\}_{n \geq 1}$,

$$P\left(\widehat{\mathbf{M}}_n \leq \mathbf{u}_n(\mathbf{x}^{(d)})\right) = P\left(\bigcap_{j=1}^d \left\{\widehat{M}_{n,j} \leq a_{n,j}x_j + b_{n,j}\right\}\right) \xrightarrow[n \rightarrow \infty]{} G(\mathbf{x}^{(d)}), \quad \mathbf{x}^{(d)} \in \mathbb{R}^d,$$

where G is a d.f. with non-degenerate margins, then Q is said to be in the max-domain of attraction of G ($Q \in D(G)$) and G is said to be a Multivariate Extreme Value (MEV) distribution function.

We will assume, without loss of generality, that the univariate marginal distributions of G are equal to F .

It is well known that the relationship between the d.f. $G(\mathbf{x}^{(d)})$, $\mathbf{x}^{(d)} \in \mathbb{R}^d$, and its marginal distributions $F(x_j)$, $j = 1, \dots, d$, can be characterized by its copula function, D_G , which exhibits a number of interesting properties, namely its stability equation:

$$D_G^t(y_1, \dots, y_d) = D_G(y_1^t, \dots, y_d^t), \quad \forall t > 0 \text{ and } (y_1, \dots, y_d) \in [0, 1]^d. \quad (1.1)$$

If the stationary sequence \mathbf{X} satisfies some mixing conditions, $D(\mathbf{u}_n(\mathbf{x}^{(d)}))$ of Hsing ([4]) or $\Delta(\mathbf{u}_n(\mathbf{x}^{(d)}))$ of Nandagopalan ([6]), and

$$P\left(\mathbf{M}_n \leq \mathbf{u}_n(\mathbf{x}^{(d)})\right) \xrightarrow[n \rightarrow \infty]{} H(\mathbf{x}^{(d)}), \quad \mathbf{x}^{(d)} \in \mathbb{R}^d,$$

where H is a d.f. with non-degenerate components, then H is also a MEV d.f.. The MEV d.f H and G can be related through the multivariate extremal index function, $\theta(\boldsymbol{\tau}^{(d)}) = \theta(\tau_1, \dots, \tau_d)$ introduced by Nandagopalan ([6]), which is a measure of clustering among the extreme values of a multivariate stationary sequence.

Definition 1.1 *A d -dimensional stationary sequence \mathbf{X} , is said to have multivariate extremal index $\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(d)}) \in [0, 1]$, if for each $\boldsymbol{\tau}^{(d)} = (\tau_1, \dots, \tau_d) \in \mathbb{R}_+^d$ there exists $\mathbf{u}_n^{(\boldsymbol{\tau}^{(d)})} =$*

$(u_{n,1}^{(\tau_1)}, \dots, u_{n,d}^{(\tau_d)}), n \geq 1$, satisfying

$$nP(X_{1,j} > u_{n,j}^{(\tau_j)}) \xrightarrow{n \rightarrow \infty} \tau_j, \quad j = 1, 2, \dots, d, \quad P\left(\widehat{\mathbf{M}}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau}^{(d)})}\right) \xrightarrow{n \rightarrow \infty} G(\boldsymbol{\tau}^{(d)}) \quad \text{and}$$

$$P\left(\mathbf{M}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau}^{(d)})}\right) \xrightarrow{n \rightarrow \infty} G^{\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(d)})}(\boldsymbol{\tau}^{(d)}).$$

As in one dimension, the extremal index is a key parameter when relating the properties of extreme values of a stationary sequence to those of independent random vectors from the same d -dimensional marginal distribution. However, unlike the one dimensional case, it is not a constant for the whole process, but instead depends on the vector $\boldsymbol{\tau}^{(d)}$.

It is now clear that the existence of $\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(d)})$ allows us to write

$$H(\mathbf{x}^{(d)}) = G^{\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(d)})}(\mathbf{x}^{(d)}) \quad \text{with} \quad \tau_j \equiv \tau_j(x_j) = -\log F(x_j), \quad j = 1, 2, \dots, d.$$

Taking $d = p + q$, it follows, as a consequence of the definition of multivariate extremal index, that the sequences $\mathbf{X}^{(p)} = \{\mathbf{X}_n^{(p)} = (X_{n,1}, \dots, X_{n,p})\}_{n \geq 1}$ and $\mathbf{X}^{(q)} = \{\mathbf{X}_n^{(q)} = (X_{n,p+1}, \dots, X_{n,p+q})\}_{n \geq 1}$ have, respectively, extremal indexes

$$\theta^{\mathbf{X}^{(p)}}(\boldsymbol{\tau}^{(p)}) = \lim_{\substack{\tau_j \rightarrow 0^+ \\ j=p+1, \dots, p+q}} \theta^{\mathbf{X}}(\boldsymbol{\tau}^{(p+q)}) \quad \text{and} \quad \theta^{\mathbf{X}^{(q)}}(\boldsymbol{\tau}^{(q)}) = \lim_{\substack{\tau_j \rightarrow 0^+ \\ j=1, \dots, p}} \theta^{\mathbf{X}}(\boldsymbol{\tau}^{(p+q)}).$$

In the notation of the extremal index we shall omit the sequence, whenever it is clear by the context and the argument of the function.

Hereinafter, let $\mathbf{Y} = (Y_1, \dots, Y_{p+q})$ and $\widehat{\mathbf{Y}} = (\widehat{Y}_1, \dots, \widehat{Y}_{p+q})$ be, respectively, two random vectors with distribution functions $G^{\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(p+q)})}$ and G , where $\mathbf{Y}^{(p)} = (Y_1, \dots, Y_p)$ and $\mathbf{Y}^{(q)} = (Y_{p+1}, \dots, Y_{p+q})$ denote two sub-vectors of \mathbf{Y} and $\widehat{\mathbf{Y}}^{(p)} = (\widehat{Y}_1, \dots, \widehat{Y}_p)$ and $\widehat{\mathbf{Y}}^{(q)} = (\widehat{Y}_{p+1}, \dots, \widehat{Y}_{p+q})$ two sub-vectors of $\widehat{\mathbf{Y}}$.

In section 2 we discuss conditions under which $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent or totally dependent. These conditions are established first by relations between the extremal indexes $\theta^{\mathbf{X}}(\boldsymbol{\tau}^{(p+q)})$, $\theta^{\mathbf{X}^{(p)}}(\boldsymbol{\tau}^{(p)})$ and $\theta^{\mathbf{X}^{(q)}}(\boldsymbol{\tau}^{(q)})$ and secondly by a coefficient that measures the strength of dependence between $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$.

The main results are illustrated in section 3 with an auto-regressive sequence and a 3-dependent sequence.

2 Main results

If the d.f. Q belongs to the domain of attraction of a MEV distribution, G , and \mathbf{X} has extremal index $\theta(\boldsymbol{\tau}^{(p+q)})$, $\boldsymbol{\tau}^{(p+q)} = (\tau_1, \dots, \tau_{p+q}) \in \mathbb{R}_+^{p+q}$, then we have

$$G^{\theta(\boldsymbol{\tau}^{(p)})}(\mathbf{x}^{(p)})G^{\theta(\boldsymbol{\tau}^{(q)})}(\mathbf{x}^{(q)}) \leq G^{\theta(\boldsymbol{\tau}^{(p+q)})}(\mathbf{x}^{(p+q)}) \leq \min\{G^{\theta(\boldsymbol{\tau}^{(p)})}(\mathbf{x}^{(p)}), G^{\theta(\boldsymbol{\tau}^{(q)})}(\mathbf{x}^{(q)})\}, \quad (2.2)$$

for each $\mathbf{x}^{(p+q)} \in \mathbb{R}^{p+q}$ and $\tau_j = -\log F(x_j)$, $j = 1, \dots, p+q$.

The inequality on the right holds true for every multivariate distribution, while the inequality on the left is a property of MEV distributions. The lower bound corresponds to the case where $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent and the upper bound corresponds to the case where $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent.

From (2.2) we obtain the following bounds for the multivariate extremal index function $\theta(\boldsymbol{\tau}^{(p+q)})$, $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$.

$$\frac{\max\{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}), \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})\}}{\gamma(\boldsymbol{\tau}^{(p+q)})} \leq \theta(\boldsymbol{\tau}^{(p+q)}) \leq \frac{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}) + \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})}{\gamma(\boldsymbol{\tau}^{(p+q)})}, \quad (2.3)$$

where

$$\gamma(\boldsymbol{\tau}^{(p+q)}) = -\log G(F^{-1}(e^{-\tau_1}), \dots, F^{-1}(e^{-\tau_{p+q}})) = \lim_{n \rightarrow \infty} nP\left(\mathbf{X}_1^{(p+q)} \not\leq u_n^{(\boldsymbol{\tau}^{(p+q)})}\right),$$

$$\gamma(\boldsymbol{\tau}^{(p)}) = \lim_{\substack{\tau_j \rightarrow 0^+ \\ j=p+1, \dots, p+q}} \gamma(\boldsymbol{\tau}^{(p+q)}) \quad \text{and} \quad \gamma(\boldsymbol{\tau}^{(q)}) = \lim_{\substack{\tau_j \rightarrow 0^+ \\ j=1, \dots, p}} \gamma(\boldsymbol{\tau}^{(p+q)}).$$

The next result follows naturally from these bounds.

Proposition 2.1 *Suppose that $Q \in D(G)$ and \mathbf{X} has extremal index $\theta(\boldsymbol{\tau}^{(p+q)})$, $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$.*

(i) *If $\widehat{\mathbf{Y}}^{(p)}$ and $\widehat{\mathbf{Y}}^{(q)}$ are independent, then $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent if and only if*

$$\theta(\boldsymbol{\tau}^{(p+q)}) = \frac{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}) + \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})}{\gamma(\boldsymbol{\tau}^{(p)}) + \gamma(\boldsymbol{\tau}^{(q)})}, \quad \boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}.$$

(ii) *If $\widehat{\mathbf{Y}}^{(p)}$ and $\widehat{\mathbf{Y}}^{(q)}$ are totally dependent, then $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent if and only if*

$$\theta(\boldsymbol{\tau}^{(p+q)}) = \frac{\max\{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}), \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})\}}{\max\{\gamma(\boldsymbol{\tau}^{(p)}), \gamma(\boldsymbol{\tau}^{(q)})\}}, \quad \boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}.$$

The necessary and sufficient conditions for $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ to be independent or totally dependent given in the previous result demand the evaluation of the extremal index function $\theta(\boldsymbol{\tau}^{(p+q)})$, in each point $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$. Nevertheless this task can be simplified with the characterizations, given in Ferreira ([3]), for independence and total dependence of the multivariate marginals of a MEV distribution. These characterizations are essential to prove the following propositions which guarantee that the independence or total dependence between $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ only depends on the value of the extremal index in some points.

Proposition 2.2 *Suppose that $Q \in D(G)$ and the sequence $\mathbf{X} = \{\mathbf{X}_n^{(p+q)}\}_{n \geq 1}$ has extremal index $\theta(\boldsymbol{\tau}^{(p+q)})$, $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$.*

The sub-vectors $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent if and only if

$$\theta(\mathbf{1}^{(p+q)}) = \frac{\theta(\mathbf{1}^{(p)})\gamma(\mathbf{1}^{(p)}) + \theta(\mathbf{1}^{(q)})\gamma(\mathbf{1}^{(q)})}{\gamma(\mathbf{1}^{(p+q)})}, \quad (2.4)$$

where $\mathbf{1}^{(k)} = (1, \dots, 1)$, $k > 1$, denotes the k -dimensional unitary vector.

Proof: Suppose that $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent. Since (2.3) holds for all $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$, we have in particular for $\boldsymbol{\tau}^{(p+q)} = (\tau, \dots, \tau) \in \mathbb{R}_+^{p+q}$, with $\tau \equiv \tau(x) = -\log F(x)$, $x \in \mathbb{R}$,

$$\theta(\boldsymbol{\tau}^{(p+q)}) = \frac{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}) + \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})}{\gamma(\boldsymbol{\tau}^{(p+q)})}.$$

Now attending to the fact that $\theta(c\boldsymbol{\tau}^{(k)}) = \theta(\boldsymbol{\tau}^{(k)})$ for each $\boldsymbol{\tau}^{(k)} \in \mathbb{R}_+^k$, $k > 1$ and $c > 0$, we can write

$$\theta(\boldsymbol{\tau}^{(p+q)}) = \theta(\mathbf{1}^{(p+q)}), \quad \theta(\boldsymbol{\tau}^{(p)}) = \theta(\mathbf{1}^{(p)}), \quad \theta(\boldsymbol{\tau}^{(q)}) = \theta(\mathbf{1}^{(q)}),$$

and from (1.1), for all $\boldsymbol{\tau}^{(p+q)} = (\tau, \dots, \tau) \in \mathbb{R}_+^{p+q}$,

$$\begin{aligned} \gamma(\boldsymbol{\tau}^{(p+q)}) &= -\log G(F^{-1}(e^{-\tau}), \dots, F^{-1}(e^{-\tau})) = -\log D_G(e^{-\tau}, \dots, e^{-\tau}) \\ &= -\log D_G^\tau(e^{-1}, \dots, e^{-1}) = \tau\gamma(\mathbf{1}^{(p+q)}), \end{aligned} \quad (2.5)$$

$\gamma(\boldsymbol{\tau}^{(p)}) = \tau\gamma(\mathbf{1}^{(p)})$ and $\gamma(\boldsymbol{\tau}^{(q)}) = \tau\gamma(\mathbf{1}^{(q)})$. Equality (2.4) is now straightforward.

On the other hand if (2.4) is verified, then for $\mathbf{x}^{(p+q)} = (x, \dots, x)$ we have

$$\begin{aligned} G_{\mathbf{Y}}(\mathbf{x}^{(p+q)}) &= G^{\theta(\mathbf{1}^{(p+q)})}(\mathbf{x}^{(p+q)}) = D_G^{\theta(\mathbf{1}^{(p+q)})}(e^{-\tau}, \dots, e^{-\tau}) \\ &= D_G^{\theta(\mathbf{1}^{(p+q)})\tau}(e^{-1}, \dots, e^{-1}) = \exp(-\tau\gamma(\mathbf{1}^{(p+q)})\theta(\mathbf{1}^{(p+q)})) \\ &= \exp(-\tau(\theta(\mathbf{1}^{(p)})\gamma(\mathbf{1}^{(p)}) + \theta(\mathbf{1}^{(q)})\gamma(\mathbf{1}^{(q)}))) = G_{\mathbf{Y}^{(p)}}(\mathbf{x}^{(p)})G_{\mathbf{Y}^{(q)}}(\mathbf{x}^{(q)}) \end{aligned}$$

and from Proposition 2.1 ([3]) we conclude that $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent. \square

Proposition 2.3 Suppose that $Q \in D(G)$, $\mathbf{X} = \{\mathbf{X}_n^{(p+q)}\}_{n \geq 1}$ has extremal index $\theta(\boldsymbol{\tau}^{(p+q)})$, $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$.

(i) If $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent then there exists $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$ with $\tau_j \equiv \tau_j(x_j) = -\log F(x_j)$, $x_j \in \mathbb{R}$, $j = 1, \dots, p+q$, such that

$$\gamma(\boldsymbol{\tau}^{(p)})\theta(\boldsymbol{\tau}^{(p)}) = \gamma(\boldsymbol{\tau}^{(q)})\theta(\boldsymbol{\tau}^{(q)}) = \theta_1\tau_1 \dots = \theta_{p+q}\tau_{p+q} = d > 0$$

and $\theta(\boldsymbol{\tau}^{(p+q)}) = \gamma\left(\frac{\boldsymbol{\tau}^{(p+q)}}{d}\right)^{-1}$.

(ii) If there exists $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$ with $\tau_j \equiv \tau_j(x_j) = -\log F(x_j)$, $x_j \in \mathbb{R}$, $j = 1, \dots, p+q$, such that

$$\gamma(\boldsymbol{\tau}^{(p+q)})\theta(\boldsymbol{\tau}^{(p+q)}) = \theta_1\tau_1 \dots = \theta_{p+q}\tau_{p+q} = d > 0,$$

then $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent.

Proof: (i) From Proposition 2.1 ([3]), if $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent, then there exists $\boldsymbol{\tau}^{(p+q)} \in \mathbb{R}_+^{p+q}$ such that

$$\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}) = \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)}) = \theta(\boldsymbol{\tau}^{(p+q)})\gamma(\boldsymbol{\tau}^{(p+q)}) = d = \theta_1\tau_1 = \dots = \theta_{p+q}\tau_{p+q},$$

with $d \in]0, 1[$. Hence

$$\begin{aligned} \theta(\boldsymbol{\tau}^{(p+q)}) &= \frac{d}{\gamma(\boldsymbol{\tau}^{(p+q)})} = \frac{d}{-\log D_G(\exp(-\tau_1), \dots, \exp(-\tau_{p+q}))} \\ &= \frac{1}{-\log D_G\left(\exp\left(-\frac{\tau_1}{d}\right), \dots, \exp\left(-\frac{\tau_{p+q}}{d}\right)\right)} = \frac{1}{\gamma\left(\frac{\boldsymbol{\tau}^{(p+q)}}{d}\right)}. \end{aligned}$$

\square

Another way to look at issues concerning independence or total dependence is to use parameters that measure the strength of dependence between $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$. We therefore define, in the following result, the dependence structure of $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ through the coefficient $\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})}$ of Ferreira ([3]). This coefficient emerged from the extremal coefficient of \mathbf{Y} , $\epsilon^{\mathbf{Y}}$, defined in Martins and Ferreira ([5]) as

$$G^{\theta(\mathbf{1}^{(p+q)})}(\mathbf{x}^{(p+q)}) = F^{\epsilon^{\mathbf{Y}}}(x), \quad x \in \mathbb{R},$$

and the relationship

$$P\left(\mathbf{Y}^{(p)} \leq \mathbf{x}^{(p)}, \mathbf{Y}^{(q)} \leq \mathbf{x}^{(q)}\right) = \left(G_{\mathbf{Y}}^{(p)}(\mathbf{x}^{(p)})G_{\mathbf{Y}}^{(q)}(\mathbf{x}^{(q)})\right)^{\frac{\epsilon^{\mathbf{Y}}}{\epsilon^{\mathbf{Y}^{(p)}} + \epsilon^{\mathbf{Y}^{(q)}}}}.$$

It is then defined as $\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = \frac{\epsilon^{\mathbf{Y}}}{\epsilon^{\mathbf{Y}^{(p)}} + \epsilon^{\mathbf{Y}^{(q)}}}$ and has the following interesting properties.

Proposition 2.4

- (i) $\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = \frac{\theta(\mathbf{1}^{(p+q)})\gamma(\mathbf{1}^{(p+q)})}{\theta(\mathbf{1}^{(p)})\gamma(\mathbf{1}^{(p)}) + \theta(\mathbf{1}^{(q)})\gamma(\mathbf{1}^{(q)})}$
- (ii) $\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = 1$ if and only if $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are independent.
- (iii) If $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent, then $\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = \frac{\max\{\epsilon^{\mathbf{Y}^{(p)}}, \epsilon^{\mathbf{Y}^{(q)}}\}}{\epsilon^{\mathbf{Y}^{(p)}} + \epsilon^{\mathbf{Y}^{(q)}}}$.

Proof: (i) Since

$$\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = \frac{\theta(\mathbf{1}^{(p+q)}) \log G(\mathbf{x}^{(p+q)})}{\theta(\mathbf{1}^{(p)}) \log G(\mathbf{x}^{(p)}) + \theta(\mathbf{1}^{(q)}) \log G(\mathbf{x}^{(q)})}$$

the result follows from (2.5).

(ii) It is an immediate consequence of (i) and Proposition 2.

(iii) If $\mathbf{Y}^{(p)}$ and $\mathbf{Y}^{(q)}$ are totally dependent then from (2.2), we have

$$\epsilon^{(\mathbf{Y}^{(p)}, \mathbf{Y}^{(q)})} = \frac{-\log \min\{G^{\theta(\mathbf{1}^{(p)})}(\mathbf{x}^{(p)}), G^{\theta(\mathbf{1}^{(q)})}(\mathbf{x}^{(q)})\}}{\theta(\mathbf{1}^{(p)}) \log G(\mathbf{x}^{(p)}) + \theta(\mathbf{1}^{(q)}) \log G(\mathbf{x}^{(q)})} = \frac{\max\{\epsilon^{\mathbf{Y}^{(p)}}, \epsilon^{\mathbf{Y}^{(q)}}\}}{\epsilon^{\mathbf{Y}^{(p)}} + \epsilon^{\mathbf{Y}^{(q)}}}.$$

□

3 Examples

Example 3.1 Let $\{Y_n\}_{n \geq 1}$ be a sequence of i.i.d. random variables with common d.f. F and consider an auto-regressive sequence of maxima $\{X_n\}_{n \geq 1}$ defined by

$$X_n = \max\{Y_n, Y_{n+1}\}, \quad n \geq 1,$$

with marginal distribution function F^2 .

Let $\{u_n^{(\tau_i)}\}_{n \geq 1}, i = 1, \dots, p$, and $\{v_n^{(\tau'_j)}\}_{n \geq 1}, j = p+1, \dots, p+q$, be sequences of real numbers such that $\lim_{n \rightarrow \infty} n(1 - F^2(u_n^{(\tau_i)})) = \tau_i$ and $\lim_{n \rightarrow \infty} nF^2(-v_n^{(\tau'_j)}) = \tau'_j$.

The sequences $\{X_n\}_{n \geq 1}$ and $\{-X_n\}_{n \geq 1}$ have, respectively, extremal indexes $\theta_1 = 1/2$ and $\theta_2 = 1$.

For sequences $\mathbf{X}_n^{(p+q)} = \begin{cases} X_{n,i} = X_n & , \quad i = 1, \dots, p \\ X_{n,i} = -X_n & , \quad i = p+1, \dots, p+q \end{cases}$, $\mathbf{X}_n^{(p)} = (X_n, \dots, X_n)$ and $\mathbf{X}_n^{(q)} = (-X_n, \dots, -X_n)$, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} P(\mathbf{M}_n^{(p)} \leq \mathbf{u}_n^{(\tau^{(p)})}) &= \exp\left(-\frac{1}{2} \max_{1 \leq j \leq p} \tau_j\right), \\ \lim_{n \rightarrow \infty} P(\widehat{\mathbf{M}}_n^{(p)} \leq \mathbf{u}_n^{(\tau^{(p)})}) &= \exp\left(-\max_{1 \leq j \leq p} \tau_j\right), \\ \lim_{n \rightarrow \infty} P\left(\mathbf{M}_n^{(q)} \leq (v_n^{(\tau'_{p+1})}, \dots, v_n^{(\tau'_{p+q})})\right) &= \exp\left(-\max_{p+1 \leq j \leq p+q} \tau'_j\right). \end{aligned}$$

Since the order statistics maximum and minimum are asymptotically independent ([2],[7]) we obtain

$$P(\mathbf{M}_n \leq (u_n^{(\tau_1)}, \dots, u_n^{(\tau_p)}, v_n^{(\tau'_{p+1})}, \dots, v_n^{(\tau'_{p+q})})) \xrightarrow{n \rightarrow \infty} \exp\left(-\frac{1}{2} \max_{1 \leq j \leq p} \tau_j - \max_{p+1 \leq j \leq p+q} \tau'_j\right)$$

and consequently $\gamma(\boldsymbol{\tau}^{(p+q)}) = \gamma(\boldsymbol{\tau}^{(p)}) + \gamma(\boldsymbol{\tau}^{(q)}) = \max_{1 \leq j \leq p} \tau_j + \max_{p+1 \leq j \leq p+q} \tau'_j$ and $\theta(\boldsymbol{\tau}^{(p+q)})\gamma(\boldsymbol{\tau}^{(p+q)}) = \frac{1}{2} \max_{1 \leq j \leq p} \tau_j + \max_{p+1 \leq j \leq p+q} \tau'_j$. Therefore

$$\theta(\boldsymbol{\tau}^{(p+q)}) = \frac{\theta(\boldsymbol{\tau}^{(p)})\gamma(\boldsymbol{\tau}^{(p)}) + \theta(\boldsymbol{\tau}^{(q)})\gamma(\boldsymbol{\tau}^{(q)})}{\gamma(\boldsymbol{\tau}^{(p)}) + \gamma(\boldsymbol{\tau}^{(q)})}.$$

Example 3.2 Let $\mathbf{U} = \{U_n\}_{n \geq 1}$ be a sequence of i.i.d. random variables with common d.f. H in the domain of attraction of the extreme value distribution F , and independent of the i.i.d. chain $\mathbf{J} = \{J_n\}_{n \geq 1}$ such that $P(J_1 = 0) = P(J_1 = 1) = 1/2$.

Let us consider a stationary 1-dependent sequence $\mathbf{Z} = \{Z_n\}_{n \geq 1}$, defined as $Z_n = U_n$ if $J_n = 0$ and $Z_n = U_{n+1}$ otherwise, and let $\mathbf{v} = \{v_n\}_{n \geq 1}$ be a sequence of normalized levels to \mathbf{Z} , and consequently also to \mathbf{U} .

We can now define a 3-dependent stationary sequence $\mathbf{X} = \{\mathbf{X}_n = (X_{n,1}, X_{n,2}, X_{n,3})\}$ as

$$(X_{n,1}, X_{n,2}, X_{n,3}) = (Z_n, Z_{n+2}, Z_{n+1}), \quad n \geq 1,$$

with common distribution function

$$T(x_1, x_2, x_3) = \frac{1}{2} \prod_{i=1}^3 H(x_i) + \frac{1}{4} H(x_1) H(\min\{x_2, x_3\}) + \frac{1}{4} H(x_2) H(\min\{x_1, x_3\})$$

belonging to the domain of attraction of

$$G(x_1, x_2, x_3) = \begin{cases} F(x_1)F(x_2)F^{\frac{1}{2}}(x_3) & , \quad x_1 < x_3 \wedge x_2 < x_3 \\ F(x_1)F^{\frac{3}{4}}(x_2)F^{\frac{3}{4}}(x_3) & , \quad x_1 < x_3 \wedge x_3 \leq x_2 \\ F^{\frac{3}{4}}(x_1)F(x_2)F^{\frac{3}{4}}(x_3) & , \quad x_3 \leq x_1 \wedge x_2 < x_3 \\ F^{\frac{3}{4}}(x_1)F^{\frac{3}{4}}(x_2)F(x_3) & , \quad x_3 \leq x_1 \wedge x_3 \leq x_2 \end{cases}$$

Now applying Proposition 2.1 of Smith and Weissman (1996) ([8]) to the sequence $U^{\mathbf{X}} = \{\max\{X_{n,1}, X_{n,2}, X_{n,3}\} = \max\{Z_n, Z_{n+1}, Z_{n+2}\}\}_{n \geq 1}$ which verifies the condition $D^{(k)}(v_n)$, $k = 2$, of Chernick et al. ([1]), we easily obtain

$$\theta^{\mathbf{X}}(1, 1, 1) = \lim_{n \rightarrow \infty} \frac{P(\max\{Z_1, Z_2, Z_3\} > v_n \geq \max\{Z_2, Z_3, Z_4\})}{P(\max\{Z_1, Z_2, Z_3\} > v_n)} = \frac{3}{10}.$$

For random vectors $\hat{\mathbf{Y}} = (\hat{Y}_1, \hat{Y}_2, \hat{Y}_3)$ and $\mathbf{Y} = (Y_1, Y_2, Y_3)$ with d.f. $G_{\hat{\mathbf{Y}}} \equiv G$ and $G_{\mathbf{Y}} \equiv G^{\theta(\tau^{(3)})}$, $\mathbf{Y}^{(2)} = (Y_1, Y_2)$ and $\mathbf{Y}^{(1)} = Y_3$ we obtain $\epsilon^{\hat{\mathbf{Y}}} = \frac{5}{2}$, $\epsilon^{\mathbf{Y}} = \frac{3}{4}$, $\epsilon^{\hat{\mathbf{Y}}^{(2)}} = 2$, $\epsilon^{\hat{\mathbf{Y}}^{(1)}} = 1$, $\epsilon^{\mathbf{Y}^{(1)}} = \frac{3}{4} = \epsilon^{\mathbf{Y}^{(2)}}$. Consequently $\epsilon^{(\mathbf{Y}^{(2)}, \mathbf{Y}^{(1)})} = \frac{1}{2}$, $\epsilon^{(\hat{\mathbf{Y}}^{(2)}, \hat{\mathbf{Y}}^{(1)})} = \frac{5}{6}$ and from Proposition 4 we conclude that neither $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ nor $\hat{\mathbf{Y}}^{(1)}$ and $\hat{\mathbf{Y}}^{(2)}$ are independent.

Nevertheless, there exists $\tau^{(3)} = (1, 1, 1)$ such that $\gamma(\tau^{(3)})\theta(\tau^{(3)}) = \theta_1\tau_1 = \theta_2\tau_2 = \theta_3\tau_3 = \frac{3}{4}$ and attending to Proposition 3 we can say that $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ are totally dependent.

Referências

- [1] Chernick, M., Hsing, T. e McCormick, W. (1991). Calculating the extremal index for a class of stationary sequences. *Advances in Applied Probability* **23**, 835-850.
- [2] Davis, R. (1982). Limit laws for the maximum and minimum of stationary sequences. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **61**, 31-42.
- [3] Ferreira, H. (2008). Measuring dependence of two multivariate extremes. *Technical report, CEAUL*, 8.
- [4] Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *Journal of Multivariate Analysis*, **29**, 274-291.
- [5] Martins, A. and Ferreira, H. (2005). Measuring the extremal dependence. *Statist. Prob. Letters*, **73**, 99-103..
- [6] Nandagopalan, S. (1990). *Multivariate extremes and estimation of the extremal index*. Ph.D. dissertation, Dept. of Statistics, University of North Carolina, Chapel Hill.
- [7] Pereira, L. (2002). *Valores Extremos Multidimensionais de Variveis Dependentes*. Tese de Doutorado, Universidade da Beira Interior.
- [8] Smith, R. L. and Weissman, I. (1996). Characterization and estimation of the multivariate extremal index. Technical report, University of North Carolina at Chapel Hill, NC, USA. In <http://www.stat.unc.edu/postscript/rs/extremal.pdf>.